

# Estimation of River Pollution Index Using Landsat Imagery over Tamsui River, Taiwan

Ying Hsuan Wang and Hong-Gyoo Sohn\*

School of Civil and Environmental Engineering, Yonsei University, Seoul 03722, Korea

Received 20 June 2018, revised 25 June 2018, accepted 25 June 2018, published online 30 June 2018

**ABSTRACT:** In-situ water quality sampling is used for accurate water quality assessment. However, in-situ water quality sampling offers limited samples and requires much time and intensive labors. Remote sensing approach has recently applied for water quality assessment. It has shown the advantage of offering a synoptic view but also more efficient and economical. In this study, we utilized Landsat Imagery to estimate the water quality of the Tamsui River basin, considered as one of the most important rivers located in the north of Taiwan. In order to monitor water quality of Tamsui River basin, a linear regression relation between the value of spectral radiance and four water quality parameters are investigated with 38 water sampling stations. Through the regression model, we could estimate river pollution index (RPI) from the predicted value of four water quality parameters. By using RPI, we can examine the pollution level of Tamsui River. The accuracy of RPI conversion of this study ranged from 32.2% to 68.2%.

**KEYWORDS:** Landsat image, River pollution index, Remote sensing, Water quality assessment

## 1. Introduction

River water is considered as a main source of fresh drinking water for human. However, the river pollution has become serious since untreated waste water was disposed to the river after overexploitation of nature resource. The most traditional method of river water quality assessment is collecting the water samples and performing laboratory analysis. This method is disadvantageous in that it cannot offer a synoptic view with limited sample because it is impossible to collect water sampling for all regions. Remote sensing has been widely used in meteorology, agriculture fisheries, disaster prevention, land management and environmental monitoring recently (Shi et al. 2009). With the advantages of efficiency, simplicity, synchronization, objectiveness and economy, remote sensing can ef-

fectively monitor river water quality in the respect of different space and time.

There are already some related studies applying remote sensing technology to estimate water quality. Wang et al. (2004) and Vazyulya et al. (2014) performed regression analysis for constructing statistical model between remote sensing image data and water quality variables. However, the remote sensing data that used in Wang et al. (2004) and Vazyulya et al. (2014) are calibrated from the in-situ radiometer measurements. The field work of in-situ radiometer measurement is not only time-consuming also impossible for performing in the whole region. In this regard, this paper aimed to develop the model between reflectance value without calibration by in-situ measurement from radiometer and water quality parameters.

\*Corresponding author: [sohn1@yonsei.ac.kr](mailto:sohn1@yonsei.ac.kr), ORCID 0000-0003-1839-3431

© Korean Society of Ecology and Infrastructure Engineering. All rights reserved.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 2. Materials and Methods

In this study, Landsat images were used to derive a proper model to estimate water quality parameters. Pre-processing of Landsat imagery includes radiometric calibration and atmospheric correction to image data. To estimate the water quality parameters from Landsat images, multiple linear regression analysis was performed by using water sampling data as ground reference data. The water sampling data were gathered and laboratory analysis was performed by Taiwan Environmental Protection Administration (EPA).

### 2.1 Study area and dataset

#### 2.1.1 Study area

Tamsui River basin (Fig. 1), which is located in the northwest of Taiwan as study area, was selected as a study area. Tamsui River, the third longest river in Taiwan, is an important source of agricultural and industrial needs and drinking. Since 1970s improper waste water disposal in Tamsui River led to rapid deterioration of water quality (Wang et al. 2014). Considering the importance of Tamsui River in water supplement, it is urgent task for monitoring water quality of Tamsui River.

#### 2.1.2 Ground reference data

Since 1976, Taiwan EPA has established a total number of 296 monitoring stations for the routine

work of river water monitoring. In the region of Tamsui River basin, there are 38 water monitoring stations (Fig. 1) that collect the water samples monthly during the past two decades. The nine water quality parameters including temperature (temp), pH, electrical conductivity (EC), dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), suspended solids (SS), ammonia (NH<sub>3</sub>-N), and coliform group are monitored monthly.

The methods of sampling and analyzing follows the standard set by Taiwan EPA for examining water and wastewater (Environmental Protection Administration, R.O.C. (Taiwan). 2010). Among those nine water quality parameters, we used four parameters, DO, BOD, SS and NH<sub>3</sub>-N for estimating river pollution index.

#### 2.1.3 Remote sensing data

The remote sensing image dataset used contains eight Landsat-5 Thematic Mapper (TM) images and two Landsat-8 Operational Land Imager (OLI) images (WRS-2 path 117 row 43) for from northern Taiwan ranging between 2000 and 2015. From analyzing the change of the water quality of different month with considering the change of weather, Landsat images from 2000 to 2015 were used to represent each month.

The Landsat images with less than 15% percent of clouds and those acquired around the same time in-situ water qualities were measured (within +/- 10 days). There was no proper image in August and December that met the above conditions. Although there were 38 in-situ water samples for each image, some of stations' measurements needed to be eliminated because of the unavoidable cloud cover on the Landsat scene.

Table 1 lists the dates and the number of the available samples of each pair of the Landsat images and in-situ water sampling data. Landsat-5 TM images sensed in seven spectral bands, which band 1 to 5 and band 7 with 30 m spatial resolution and band 6 with 120 m spatial resolution, simultaneously. On the other hand, Landsat-8 OLI images sensed in nine spectral bands, which band 1 to 7 and band 9 with 30 m spatial resolution and band 8 with 15 m spatial resolution. Since



**Fig. 1.** Tamsui river basin and 38 water sampling stations established by Taiwan EPA.

**Table 1.** The dates and the number of the available samples of each pair of the Landsat images and in situ water sampling data

Image Date (yyyy-mm-dd)	In-situ Date (yyyy-mm-dd)	No. of available station	Image Date (yyyy-mm-dd)	In-situ Date (yyyy-mm-dd)	No. of available station
2002-01-12	2002-01-21	33	2015-06-09	2015-06-04	30
2011-02-06	2011-02-08	26	2006-07-02	2006-07-03	27
2007-02-27	2007-03-01	28	2006-09-04	2006-09-04	31
2015-04-06	2015-04-09	29	2000-10-05	2000-10-06	25
2009-05-07	2009-05-07	22	2004-11-01	2004-11-03	26

**Table 2.** Band selection of Landsat-5 TM and Landsat-8 OLI in this study

Landsat-5 TM		Landsat-8 OLI	
Spectral Band	Wavelength (μm)	Spectral Band	Wavelength (μm)
Band 1 - Blue	0.450 - 0.520	Band 2 - Blue	0.450 - 0.515
Band 2 - Green	0.520 - 0.600	Band 3 - Green	0.525 - 0.600
Band 3 - Red	0.630 - 0.690	Band 4 - Red	0.630 - 0.680
Band 4 - Near Infrared	0.760 - 0.900	Band 5 - Near Infrared	0.845 - 0.885
Band 5 - Near Infrared	1.550 - 1.750	Band 6 - Short Wavelength Infrared	1.560 - 1.660
Band 7 - Mid-Infrared	2.080 - 2.350	Band 7 - Short Wavelength Infrared	2.100 - 2.300

both Landsat-5 and Landsat-8 images were used for construct the model for estimating water quality parameters, this study used the bands at 30 m spatial resolution, available in both Landsat-5 TM and Landsat-8 OLI sensors. The bands of each sensor used for the analysis are listed in Table 2.

## 2.2 Method

### 2.2.1 Landsat image pre-processing

Landsat image pre-processing included radiometric calibration and atmospheric correction. The data of Landsat-5 TM sensors is in the form of 8-bit digital number between 0 to 255, the numbers need to be converted to top of atmosphere (ToA) reflectance by two-step processes. First, the digital numbers (DNs) value need to be converted into radiance by Eq. 1,

$$L_{\lambda} = gain * DN + bias \quad (Eq. 1)$$

where  $L_{\lambda}$  is the radiance value of the cell ( $W m^{-2} sr^{-1} \mu m^{-1}$ ), DN is the digital numbers value of the cell, gain is the gain value of specific band and bias is the bias value of

specific band. After conversion to radiance, the radiance needed to be converted to ToA reflectance by Eq. 2,

$$\rho_{\lambda} = \pi * L_{\lambda} * \frac{d^2}{ESUN_{\lambda}} * \cos \theta_s \quad (Eq. 2)$$

where  $\rho_{\lambda}$  is the unit-less reflectance value,  $L_{\lambda}$  is spectral radiance of the cell from the previous step,  $d$  is earth-sun distance in astronomical units,  $ESUN_{\lambda}$  is mean solar exoatmospheric irradiance and  $\theta_s$  is solar zenith angle.

Meanwhile, Landsat-8 OLI sensor images DNs value is 16-bit integer value (Landsat 8 Data User Handbook, year). These values can be convert to ToA reflectance by Eq. 3,

$$\rho_{\lambda}' = M_{\rho} * Q_{cal} + A_{\rho} \quad (Eq. 3)$$

where  $\rho_{\lambda}'$  is the unit-less reflectance value,  $M_{\rho}$  is reflectance multiplicative scaling factor for specific band,  $A_{\rho}$  is reflectance additive scaling factor for specific band and  $Q_{cal}$  is the digital value of the cell.

After the DNs were converted to the reflectance values, atmospheric correction was performed for removing the effects of atmosphere from water vapor, aerosol, etc.

### 2.2.2 Multiple linear regression analysis

This study performed multiple linear regression analysis for relating the reflectance value of each bands in Landsat images to the actual measurement of four water quality parameters included DO, BOD, NH<sub>3</sub>-N and SS from in-situ water samples. Through the GPS location of water sampling stations, the reflectance value near the station was acquired. Then, the regression model of each water quality parameters for every month was established. The extracted reflectance values of each bands were used as independent variables and the water quality parameters were used as dependent variables to perform multiple linear regression analysis.

### 2.2.3 River pollution index conversion

RPI has been used by the Taiwan EPA to evaluate the conditions of surfaces water pollution. RPI is a simple integrated indicator for determining the level of pollution of a river. The value of RPI is calculated by 4 parameters in water quality including DO, BOD, SS and NH<sub>3</sub>-N. Each parameter is divided into four ranks and score was assigned. The score conversion of each water parameter are shown in Table 3. The RPI is calculated by Eq. 4:

$$RPI = \frac{1}{4} \sum_{i=1}^4 P_i \quad (\text{Eq. 4})$$

## 3. Result and Discussion

### 3.1 Multiple linear regression model

To determine the relationship between the reflectance from Landsat images and water quality parameters, multiple linear regression models were established. Water quality parameters included DO, BOD, SS and NH<sub>3</sub>-N. Regression models showed the results with R square value ranging 22.2% to 71.5% in DO, 13.7% to 46.5% in BOD, 21.6% to 64.0% in SS and 9.0% to 57.8% in NH<sub>3</sub>-N from ten Landsat images with represent of different month. The R squares of four water parameters were shown in Fig. 2. Fig. 2 shows that the model of May has the best performance and the model of February has the worst performance in four concentrated water quality parameters.

### 3.2 Estimation of river pollution index

Based on the constructed models, the prediction value of four water quality parameters was determined. The predication value can be converted to RPI value and reveal the level of pollution of the location through the RPI value. After converting to the level of pollution by RPI, the percentage of predicting the level of pollution correctly ranged from 32.2% in September to 68.2% in May. The RPI conversion results from the remote sensing image of May and September are shown in Fig. 3.

**Table 3.** River pollution index

Items	Ranks			
	Unpolluted	Negligibly Polluted	Moderately Polluted	Severely Polluted
DO (mg/L)	Above 6.5	4.6-6.5	2.0-4.5	Under 2.0
BOD (mg/L)	Under 3.0	3.0-4.9	5.0-15	Above 15
SS (mg/L)	Under 20	20-49	50-100	Above 100
NH <sub>3</sub> -N (mg/L)	Under 0.5	0.5-0.99	1.0-3.0	Above 3.0
Index Score (Pi)	1	3	6	10
RPI	Under 2	2.0-3.0	3.1-6.0	Above 6.0

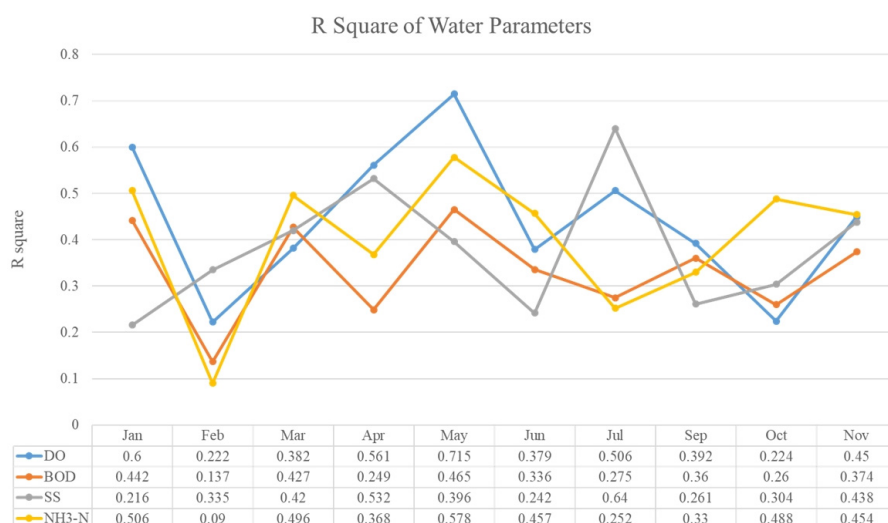


Fig. 2. R square of four water parameters in each month.

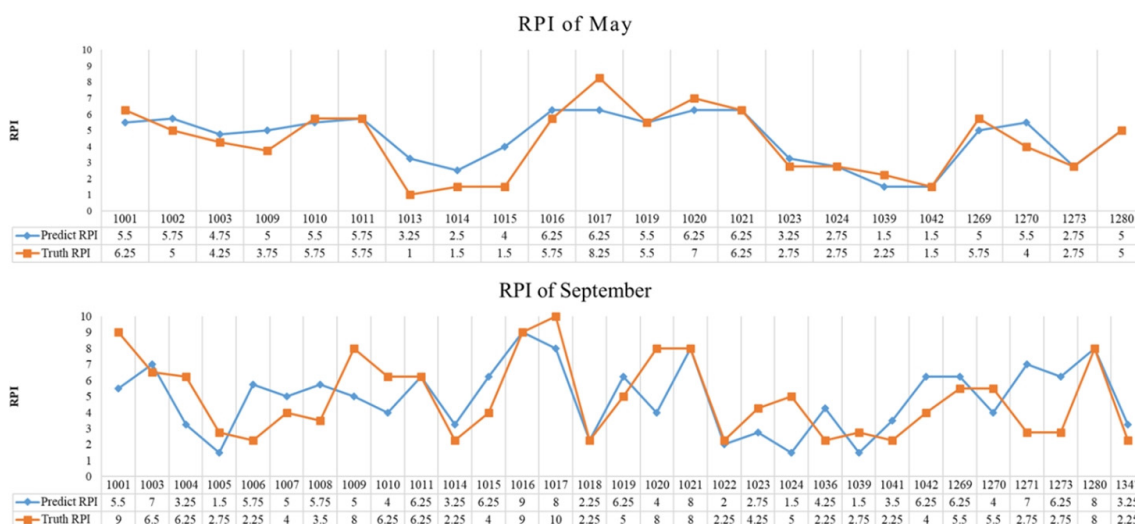


Fig. 3. Conversion result of RPI in May and September.

## 4. Conclusion

The possibility of RPI estimation from remote sensing data is shown in this study. The multiple linear regression model was built with the concentrations of four water quality parameters (DO, BOD, SS and NH<sub>3</sub>-N) respectively. The R square value is ranging from 22.2% to 71.5% in DO, 13.7% to 46.5% in BOD, 21.6% to 64.0% in SS and 9.0% to 57.8% in NH<sub>3</sub>-N. Through the regression analysis, this study confirmed that the accuracy of RPI conversion which ranges

from 32.2% to 68.2% is moderate. With the result from this study, the following implications could be drawn. First, the remote sensing images including Landsat 5 and Landsat 8 images needed to be done further correction since there must be some difference between two satellite. Second, although remote sensing technology can effectively monitor the river water pollution, still, more in-situ water samples are required for developing a better statistic model. Since there might be a bias when the sample size is small. If the further correction between different satellite and more

samplings are available, we can look forward to retrieved a better model for water quality parameters estimation. Also, the improvement of RPI estimation can be achieved with an improved water quality parameters models.

## References

- Environmental Protection Administration, R.O.C. (Taiwan). 2010. Environmental Water Quality Information, <https://wq.epa.gov.tw/Code/Business/ItemMethod.aspx>. Accessed 25 November 2017.
- Shi, M.L., Yang, Z.G., and Yan, K.H. 2009. Using unsupervised method to classify the space of river pollution in the image of the satellites. *Journal of Photogrammetry and Remote Sensing* 14(4): 287-302.
- Singhvi, B.S., Singh, B.G., and RP, A. 2015. water quality assessment for open water bodies by remote sensing: A review. *International Journal of Engineering Technology and Computer Research* 3(3).
- Vazyulya, S., Khrapko, A., Kopelevich, O., Burenkov, V., Eremina, T., and Isaev, A. 2014. Regional algorithms for the estimation of chlorophyll and suspended matter concentration in the Gulf of Finland from MODIS-Aqua satellite data. *Oceanologia* 56(4): 737-756.
- Wang, Y.B., Liu, C.W., Liao, P.Y., and Lee, J.J. 2014. Spatial pattern assessment of river water quality: implications of reducing the number of monitoring stations and chemical parameters. *Environmental Monitoring and Assessment* 186: 1781-1792.
- Wang, Y., Xia, H., Fu, J., and Sheng, G. 2004. Water quality change in reservoirs of Shenzhen, China: detection using LANDSAT/TM data. *Science of the Total Environment* 328(1-3): 195-206.